

DOCUMENT RESUME

ED 022 631

RE 001 309

By-Powell, William R.

REAPPRAISING THE CRITERIA FOR INTERPRETING INFORMAL INVENTORIES.

Pub Date 25 Apr 68

Note-18p.; Paper presented at International Reading Association Conference, Boston, Mass., April 24-27, 1968.

EDRS Price MF-\$0.25 HC-\$0.80

Descriptors-EDUCATIONAL TESTING, *ELEMENTARY SCHOOL STUDENTS, *INFORMAL READING INVENTORY, LANGUAGE LEARNING LEVELS, *READING COMPREHENSION, READING LEVEL, *READING RESEARCH, READING SKILLS, READING TESTS, *WORD RECOGNITION

The widely accepted 95 percent word recognition criterion used for designating the instructional reading level is re-evaluated. A survey of the original study of Betts and Killgallon indicated (1) that the 95 percent word recognition criterion was based on a 41-case sample at the fourth-grade level, (2) that repetition errors occurred most frequently and were counted as scoreable errors, and (3) that the child read the paragraph silently first, then orally. A sample of 178 average-achieving pupils in grades 1 to 6 was drawn. The highest reading level with a comprehension score nearest the 70 percent cutoff level was determined for each subject. The lowest percent of word recognition accuracy within the limits set by the comprehension score (70 percent) was recorded. Mean scores were computed for each grade level and in various combinations. The data indicated that pupils in grades 1 and 2 could tolerate on the average an 85 percent word recognition score and still maintain 70 percent comprehension. Pupils in grades 3 through 6 could tolerate on the average 91 to 94 percent word recognition score. It is suggested that the currently used informal inventories be viewed more as a methodology with guidelines rather than as a test instrument. A table and references are included. (CM)

William R. Powell
College of Education
University of Illinois
Urbana, Illinois 61801

U. S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Session 2C - Evaluation
Thursday, April 25, 3:00-4:00 p.m.
International Reading Association
Boston, Massachusetts
1968

REAPPRAISING THE CRITERIA FOR INTERPRETING
INFORMAL INVENTORIES

RE 001
309
ED 022631
In discussing the basic content of this presentation with a few strong adherents of the present system of interpreting informal inventories, one reply made in a skeptical tone was, "No one will believe you." Knowing this is the way many of you will feel about the central issue of my remarks, I view this presentation with some ambivalence. It is not considered fashionable to attempt to reshape a widely recognized "clay idol."

Another remark offered in support of the way we now interpret the informal was confidently given, "I know it is correct because it works." My immediate thought reaction was, 'My car works, but it is badly out of tune.' This

was a natural reaction on my part, since I had just driven several miles to our place of meeting and was wondering if my car would get me home again without incident.

Nevertheless, these remarks were troublesome, because they gave the implication that we should not explore our present standards and concepts to gain greater clarification--in other words, "Don't rock the boat!"

However, in my own mind, all the pieces to the puzzle of the informal do not fit neatly into place. For some three years now, I have had the growing feeling our present criterion of 95% word perception accuracy and 75% comprehension for determining the instructional reading level was in error. Adhering to the philosophic tenet that to err is to sin, I offer the following comments as a petition for our waywardness.

Lest you think I have completely "flipped," let me state my position clearly at the outset that I firmly believe in the effective use of informal reading inventories. For my own purposes, I use the informal, teach it, and urge its use.

Informal reading inventories provide information pertinent to the assessment of reading ability to a degree far surpassing the quality and quantity of evaluation obtainable with standardized instruments. It is not the intent of this paper to launch an attack on the informal;

only to suggest modification where it is indicated. It is believed that the reported observations that follow will make the informal more useful and accurate than it is presently.

The present use of the informal inventory, as effective as it is, does not justify the continuance of error in its criteria, nor the discouragement of research to increase its effectiveness. As a matter of fact, there is a dearth of research supporting any system of criteria for interpreting informal reading inventories. A wealth of opinion and intuition does exist regarding criterion levels for evaluating reading by this technique, but little valid research data to support those suppositions.

The central premise of this paper is that the word-recognition criterion used for designating the instructional reading level which has typically been applied for some twenty-five years is incorrect. An inspection of the minimal data available would suggest that the criterion was incorrectly formulated initially, and therefore, has been incorrectly applied and perpetuated all this time.

With the premise stated, what follows in the remainder of this paper is the presentation of the data and information which would support that position, suggest a probable correct percentage zone for the word-recognition criterion, and raise questions about the informal inventory with implications for

further research.

Since its origin, the informal reading inventory has undergone several modifications, both in terms of administrative procedure and of the percentages used to designate reading levels. Reading authorities have varied in opinion regarding the exact percentage at the instructional reading level. (Variance also occurs on the other levels, but this paper will focus only on the instructional level.) Percentages suggested for the lower limits of needed accuracy in word-recognition and in comprehension are as follows: for example, with word-recognition first and comprehension second in each instance, E. A. Betts (and P. A. Killgallon), 95 and 75; A. J. Harris, 95 and 75; M. Botel, 95 and 75; R. A. Kress and M. S. Johnson, 95 and 75; M. C. Austin and M. H. Huebner, 95 and 75; L. R. Wheeler and E. H. Smith, 95 and 70; R. Karlin, 90 and 70; N. B. Smith, 80 and 70; and L. Cooper, 98 and 70 (primary), and 96 and 60 (intermediate). With few exceptions, even in those instances where the criteria have been modified, the Betts-Killgallon paradigm has been generally retained.

Betts-Killgallon Criteria

Betts (1), who originated the informal reading inventory, defines the criteria on the basis of an investigation by P. A. Killgallon (4). Killgallon, working under Betts' direction, completed the study which appears to have led

directly to the development of the criteria for the informal inventory. The major concern of the Killgallon thesis, however, was not the production of criteria for the informal. The criteria appear to have emerged as a kind of by-product of the investigation.

Initially, Killgallon arbitrarily established the lower-limits for the instructional level as at least 50 per cent comprehension or better, and a word-recognition ratio of one error in every fourteen running words, i.e., 92.86 per cent. These limits automatically excluded any exploration of findings outside the zero to seven per cent range for word-recognition errors. Even though his hypothesis was suggestive of the 93 per cent level for word-recognition accuracy, Killgallon shifted, without apparent justification, to the 95 per cent level in word-recognition and 75 per cent level for comprehension in his conclusions. This shift in percentage was based on the data gathered from forty-one fourth grade protocols.

An examination of, and recomputation of, Killgallon's percentages of those forty-one fourth grade cases provided in the appendix of his study almost perfectly support his hypothesis, not his conclusions. Of the 39 cases which had an instructional level (two had no observable instructional level) when holding comprehension constant at 75 per

cent or higher, the mean word-recognition per cent of Killgallon's subjects was 93.8 per cent, or approximately one error in every sixteen running words. This percentage figure does not, however, take into account that repetitions were counted as scoreable errors, and this type of error occurred the most frequently.

When one realizes the fact that the criteria for the informal reading inventory as commonly used and promulgated are based on only 41 cases at one grade level (fourth) and that the data and the general conclusions of the basic study for its definition are not congruent, then serious questions regarding the accuracy of the instrument become apparent, and re-evaluation is in order. Whether a few percentage points in the word-recognition score will make a difference in determining the instructional level, only further research will reveal.

Cooper Investigation

Cooper (2), conducted the only experimental investigation for the establishment of criteria for the informal reading inventory. He was the first to suggest an alteration of the original Betts criteria and the only individual to support his modifications with objective evidence. The Cooper criteria are even more stringent than the original Betts standards. He indicated that the primary level should have a word-recognition score of at least 98 per

cent supported by at least 70 per cent comprehension score, while the intermediate levels require 96 per cent and 60 per cent, respectively.

The Cooper effort is commendable in its intent, but unfortunately, the findings of his study are weakened by certain methodological difficulties, such as, examiner and material variance, lack of adequate range for reading responses, and inconsistent critical ratio patterns. The only certain conclusion from his data would seem to support the conception that children in frustration reading material do not grow in reading skill, as measured by a standardized test, at a rate commensurate with those children who are not in frustration reading material. However, with procedural modification, the Cooper study could provide a model for direction in a further investigation in the search for more accurate criteria.

Daniels and Schummers Reports

Daniels (3) used a modified criterion which was suggested by Mazurkiewics and which permitted more latitude in the word-recognition score. The modified standards allowed for the designation of the instructional level a word-recognition score of 90 to 97 per cent, with an accompanying comprehension score of 70 to 79 per cent. Daniels indicated that the modified scores, with greater error range, yielded almost the same results and correlations as the

original Betts criteria. He suggested the modified scoring procedure has no advantage over the original standards, and he may well be correct, but further research would be necessary to substantiate this impression.

In a study conducted by Schummers (5), one of the subsidiary findings revealed that third grade children rarely read with 95 per cent accuracy in sight oral reading in material of their grade level or even one year lower than their actual grade placement. Schummers advised caution in applying standard oral word-recognition accuracy indices as a means of estimating reading levels. According to his data, the 95 per cent level of accuracy in oral reading at sight appeared to be too high, even for the most able readers in a third grade group.

Although Schummers does not specifically indicate an ^{ex} indice level for reading placement, by the examination of his data and by applying simple division to figures in Table 16 and Table 21, it becomes apparent that his third grade random sample of 237 pupils were making approximately seven scoreable errors per one hundred running words. This figure, of course, closely resembles the actual data figures and original hypothesis of Killgallon.

One other point is essential to the proper comparative interpretation of the studies discussed. This point centers on the awareness of and understanding of the procedures

used by each investigator in administering the informal. In the Cooper, Daniels, and Schummers studies, the oral portion of the informal was administered to the pupils with no prior exposure of the material, i.e., oral reading at sight. This is the commonly accepted practice today. However, the administrative procedures used in the original study with the informal differ from this practice. The Betts-Killgallon approach of the 1940's was to have the child read the paragraph silently first, then orally, followed by the comprehension questions. The oral reading, after preparation, was used for determining the word-recognition score. A second paragraph at the same level was read silently with subsequent comprehension check for the silent portion of the informal. This pattern follows the model for a directed reading lesson. This original approach, without much doubt, would have the effect of raising the percentage of accuracy of the word-recognition score. Coupled with the fact that the original data as re-evaluated revealed an average instructional level percentage of 93.8, when the procedure used had the silent precede the oral reading, then the natural inference would be that the Killgallon 93.8 figure is not only spurious, ^{how much lower} but high for the population studied. ^h The word-recognition score would be under an oral at sight approach. ~~The per cent~~ is a matter for speculation.

To sum up to this point in the discussion, the evidence available would suggest that the 95 per cent word-recognition criterion for determining the instructional reading level is not experimentally supported. Further, the data would tend to indicate that a more accurate standard for judging oral sight reading might be discovered and established at a lower level than the 95 per cent measure.

Powell Investigation

To test the hypothesis that the word-recognition criterion was lower than the 95 per cent level and to attempt to pinpoint the probable level, I examined the informal profiles for each of 178 protocols involving average-ability, average-achieving (by standardized measures), middle-class youngsters. The sample was composed of 31, 28, 32, 32, 22, 33 pupils in grades one through six, respectively. All the data had been carefully collected by Dr. Peggy Ransom, Dr. Delbert Patty, and Dr. Clara Kirby.

Initial observation of the total cases revealed that the two components under inspection, i.e., word-recognition and comprehension, were not atypically skewed to the advantage of either variable. Any abnormally imbalanced profile would probably be accounted for by individual variation, rather than a total improper instructional

emphasis of the school program.

The procedure used in evaluating each of the 178 cases was to hold the comprehension score constant at 70 per cent or higher. By this procedure, the first criterion to be inspected and evaluated was the comprehension score column of each profile. The lowest acceptable score for classification at the instructional level was for the subject to have at least a comprehension percentage of 70 or higher. The reason the 75 per cent comprehension criterion score was not used was to mitigate the effects of the comprehension score which could have been influenced by the number of questions asked of the subject. Observation revealed that this precaution was not truly necessary, as only a very small number of the cases would have been so affected.

Each protocol was first examined to locate the highest reading level with a comprehension score nearest the 70 per cent criterion, but still higher than the arbitrary cut-off level. This level determined the point of entry into the word-recognition score column. The scores in the word-recognition column up to and at that level were scanned to determine the lowest per cent of word-recognition accuracy within the limits set by the comprehension score.

The rationale used for this procedure was that if the youngster's comprehension per cent remained continuously

at an acceptable level, then the fluctuation in word pronunciation errors was tolerable to the reader. This lowest percentage figure was recorded for each case as indicative of the possible tolerable word-recognition score for that child, and this figure was the number used in the subsequent computations.

The results of the mean word-recognition scores are provided in Table 1. Mean scores were computed for each grade level separately and in various combinations: primary; intermediate; grades one and two; grades three and four; grades five and six.

Insert Table 1 about here

The findings were indeed surprising--to say the least. While the lower per cents or lower word-recognition ratios were anticipated and supportive of the hypothesis, the degree of difference was totally unexpected.

The data suggest that the mean scores of grades one and two tend to cluster together percentage-wise, and the mean percentages of grades three, four, five, and six form a relatively similar percentage zone. The data clearly indicate that pupils in grades one and two could tolerate on the average an 85 per cent word-recognition score and still maintain seventy per cent comprehension. To say I

William R. Powell
University of Illinois
IRA, Boston, 1968

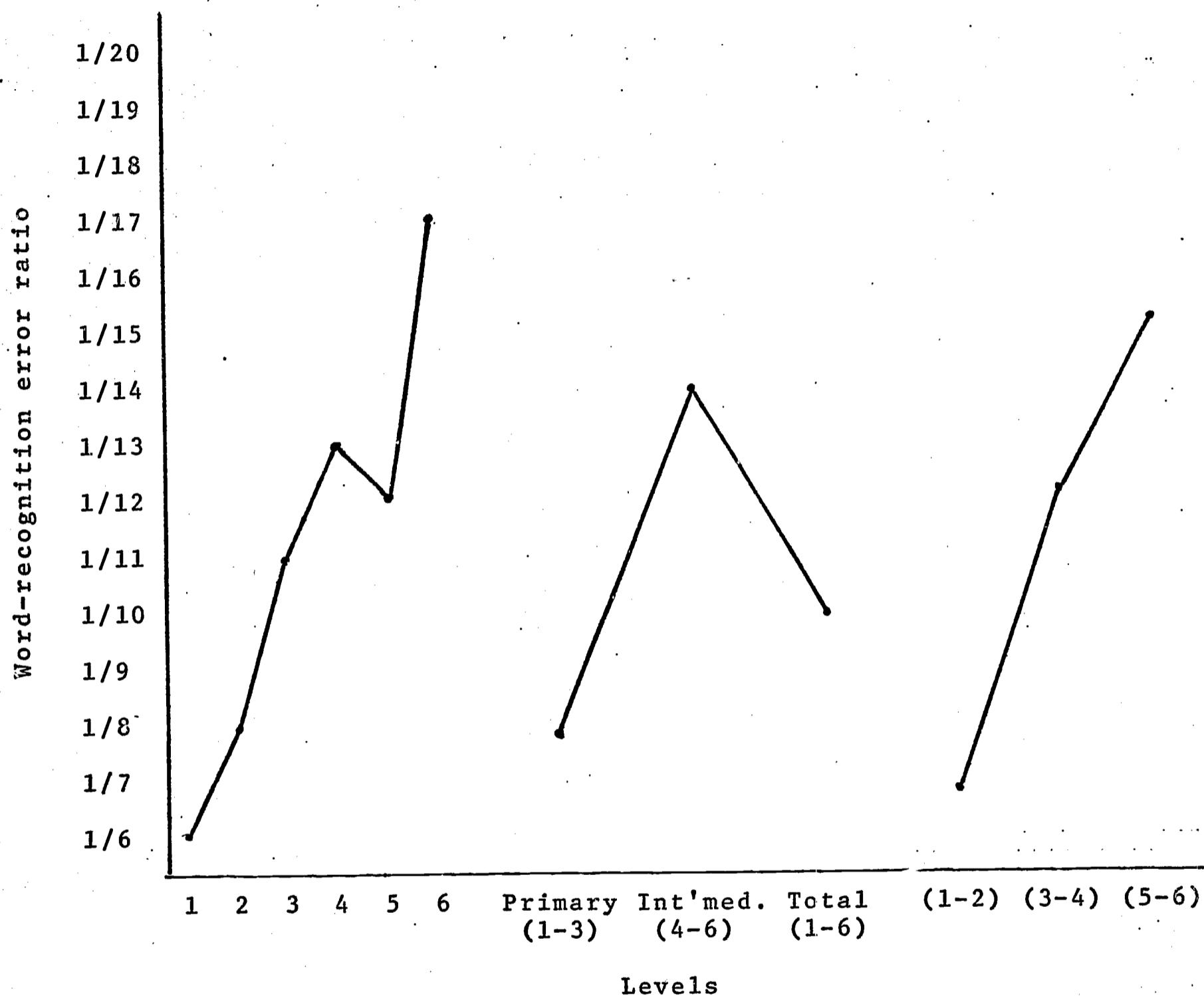


Table 1. Average word-recognition error ratio at each grade level and grade clusters.

found this astounding is an understatement! The finding that pupils in grades three through six could tolerate on the average a 91 to 94 per cent word-recognition score while maintaining 70 per cent comprehension is commensurate with the data of Killgallon and Schummers. This finding was expected.

At any level, the 95 per cent model as suggested by Betts was not supported. Of the 178 cases, slightly over one-half of the cases would not have met the 95 per cent word-recognition criterion when comprehension was the point of entry into the individual profile.

So dubious was I upon the completion of the computations that I repeated the process twice. No computational error was found. Still skeptical, I spent a long evening pulling cases out of the clinic files in the Center for Reading Research and Instruction, so more cases could be inspected. I knew the clinic population was atypical, but it was my hope they would at least verify or disprove the data I had. After locating in the clinic files twelve cases at each grade level, two through six, I followed the same procedure with these cases as described earlier. Surprisingly, these cases tended to support the previous findings. Therefore, I offer the findings as probably tenable.

Discussion

The evidence presented would tend to indicate that the widely accepted criterion of 95 per cent word-recognition score for the instructional reading level as determined through the informal inventory needs a complete reappraisal. The data, past and present, would suggest that the 95 per cent standard is too high for the majority of the youngsters in grades one through six.

Further, the data suggest that the younger child can tolerate more word-recognition error and maintain an acceptable comprehension level than youngsters in grades three through six. Whether this difference is due to the complexity of the language used for reading between these two groups, the difference in the depths of concepts presented in the reading materials at the upper levels, both language and concepts, or other factors not immediately discernible can only be verified through further research.

Excluding the Cooper study, first and second grade pupils have not been investigated in any attempt to establish criteria for those levels. Evidently the criterion has been generalized to apply to those lower grade levels. The Killgallon, Daniels, and Schummers studies were conducted on fourth, third, and third grade students, respectively. Therefore, perhaps the 85 per cent word-recognition score for the first two grades is not so astonishing. There is no base for comparison.

However, the study by Cooper did include second grade pupils. Also, Cooper indicates that the criterion for the lower grades should be more stringent, i.e., a higher word-recognition per cent, than for the intermediate grade pupils. The evidence in the data above indicates a trend just the reverse of the Cooper conclusion. Even though it was pointed out that the Cooper study contained some procedural errors, the proper action called for now is new experimental data to test the directional contradictions.

Quite possibly the answer may have been suggested by the Daniels investigation. He implies that the per cent of word-recognition error, within a given range, is not an important variable in determining the instructional reading level. The limits of this range are yet to be determined; and since the Daniels study used third grade subjects and his modified range (90-97) included the average found for third-grade pupils above (90.84), any conclusion is risky.

Summary

In spite of questions yet unresolved, there is strong evidence to safely conclude that the widely used Betts criterion of word-recognition for determining the instructional reading level through the informal inventory should be held suspect. Betts originated an excellent device for evaluating reading, and perhaps if we viewed the informal more as

a methodology with defined guide lines and less as a test instrument, attempts would constantly be made to perfect those guide lines.

REFERENCES

1. Betts, Emmett A. Foundations of Reading Instruction. New York: American Book Co., 1950.
. "Reading Problems at the Intermediate Grade Level," Elementary School Journal, 40 (June, 1940), 737-746.
. "What Are the Reading Levels of Your Pupils?" The School, 34 (Nov., 1945), 194-195.
2. Cooper, J. Louis. "The Effect of Adjustment of Basal Reading Materials on Reading Achievement." Unpublished Doctoral Dissertation, Boston University, Boston, 1952.
3. Daniels, John E. "The Effectiveness of Various Procedures in Reading Level Placement," Elementary English, 39 (Oct., 1962), 590-600.
4. Killgallon, Patsy A. "A Study of Relationships Among Certain Pupil Adjustments in Language Situations." Unpublished Doctoral Dissertation, The Pennsylvania State University, University Park, 1942.
5. Schumers, John L. "Word Pronunciation in the Oral Sight-Reading of Third Grade Children." Unpublished Doctoral Dissertation, University of Minnesota, Minneapolis, 1956.